REPPERGER RESEARCH INTERN PROGRAM

RESEARCH PROJECT #: AFRL-RHW-23-11

The Impact of Bias on AI Development

PROJECT DESCRIPTION: Artificial intelligence (AI) systems and machine learning (ML) algorithms have streamlined decision making in our daily lives: from the ordinary, such as algorithms suggesting a movie on a streaming platform, to weightier topics, like who is automatically approved or rejected for a bank loan (Mehrabi et al., 2021). AI is also now embedded in cognitive warfare as nations attempt to influence the sixth operational, or Human (Cognitive/Brain), domain (Masakowski, 2022). However, there can be negative long-term consequences when we assume that AI systems are synonymous with "unbiased" or fail to study possible vulnerabilities in these systems. Although AI research is examining how to use computational approaches to mitigate bias after an algorithm is designed (Chen & Jou, 2021), less is known about how to study possible biases during the early stages of development, such as during the data annotation process. Implicit bias, often measured with timed responses, (Greenwald et al., 1998) is defined as attitudes that are relatively more unintentional and uncontrollable compared to more explicit (or self-reported) bias, which is more controlled and intentional (Gawronski & De Houwer, 2013). This project will focus on what types of cognitive (e.g., confirmation bias; framing bias; loss aversion) and social biases (e.g., sexism; ethnocentrism) can affect algorithm development, how to best measure possible bias (implicit and explicit measures), and possible mitigation strategies. Interns will actively participate in the research process for ongoing projects, including spin-off study idea generation for an independent project. Soft skill development will include discussing research ideas, design, and results with scientists in their primary field of study and from other disciplines; hard skill development will focus on building experience using R for data analysis and using social media analysis tools to collect data trends.

ACADEMIC LEVEL: Bachelor's, Master's, PhD

DISCIPLINE NEEDED:

- Social Psychology
- Experimental Psychology
- Psychology

RESEARCH LOCATION: Virtual or In-Person at Wright-Patterson Air Force Base, Dayton, Ohio

RESEARCH MENTOR: Christine Vitiello, PhD
Experimental Psychology, University of Florida, 2021

Dr. Christine Vitiello is a Research Psychologist for the Air Force Research Laboratory's Mission Analytics Branch (711/ RHWAR). She is a social psychologist interested in how context, attitudes, and individual differences affect online and offline behavior and the impact of cognitive and social biases on Artificial Intelligence (AI) design and application. *Photo courtesy of the U.S. Air Force Research Laboratory.*